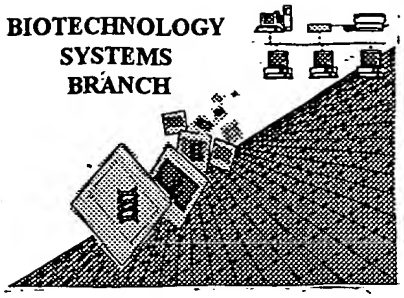
7510

1645

1600

RAW SEQUENCE LISTING ERROR REPORT



RECEIVED
JAN 2 2 2002
TECH CENTER 1600/2900

RECEIVED
APR 0 8 2002
TECH CENTER 1600/2900

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/623,568
Source: 1600
Date Processed by STIC: 1/15/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom, including:

1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm**> , **EFS Submission User Manual** - ePAVE)
2. U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/623,568

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFT

1	Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2	Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3	Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4	Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5	Variable Length	Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6	PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7	Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading). (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8	Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9	Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10	Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11	Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12	PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13	Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

RECEIVED
APR 0 8 2002
TECH CENTER 1600/2900

1600

RAW SEQUENCE LISTING DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

pp 1-2
Does Not Comply
Corrected Diskette Needed

3 <110> APPLICANT: Miller, Barbara
4 Osmani, Stephen
5 Clawson, Gary
6 Zhang, Min-Ying
7 Norris, James
9 <120> TITLE OF INVENTION: Use of Human Homolog Of A Nuclear Migration Gene For Treatment And
10 Diagnosis Of Cancer
12 <130> FILE REFERENCE: PSU-0016
14 <140> CURRENT APPLICATION NUMBER: 09/623,568
15 <141> CURRENT FILING DATE: 2001-03-23
17 <150> PRIOR APPLICATION NUMBER: 60/076,885
18 <151> PRIOR FILING DATE: 1998-03-05
20 <150> PRIOR APPLICATION NUMBER: PCT US99/04996
21 <151> PRIOR FILING DATE: 1999-03-05
23 <160> NUMBER OF SEQ ID NOS: 16
25 <170> SOFTWARE: PatentIn version 3.1
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 14
29 <212> TYPE: PRT
30 <213> ORGANISM: artificial Sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Peptide
35 <400> SEQUENCE: 1
37 Gly Cys Met Val Glu Lys Met Met Tyr Asp Gln Arg Gln Lys
38 1 5 10
41 <210> SEQ ID NO: 2
42 <211> LENGTH: 15
43 <212> TYPE: PRT
44 <213> ORGANISM: artificial Sequence
46 <220> FEATURE:
47 <223> OTHER INFORMATION: Peptide
49 <400> SEQUENCE: 2
51 Asn Gly Ser Leu Asp Ser Pro Gly Lys Gln Asp Thr Glu Glu Asp
52 1 5 10 15
55 <210> SEQ ID NO: 3
56 <211> LENGTH: 24
57 <212> TYPE: DNA
58 <213> ORGANISM: artificial Sequence
60 <220> FEATURE:
61 <223> OTHER INFORMATION: Oligonucleotide
63 <400> SEQUENCE: 3
64 ttctgttcgt ctgaagttgg cagc 24
67 <210> SEQ ID NO: 4
68 <211> LENGTH: 24
69 <212> TYPE: DNA
70 <213> ORGANISM: artificial Sequence
72 <220> FEATURE:

insufficient explanation - give source of genetic material (see item 11 on Error Summary Sheet)

same error

same type of error

RAW SEQUENCE LISTING DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

```
 73 <223> OTHER INFORMATION: Oligonucleotide
 75 <400> SEQUENCE: 4
 76 caatgaagtg aaggtggagg agag                                          24
 79 <210> SEQ ID NO: 5
 80 <211> LENGTH: 38
 81 <212> TYPE: DNA
 82 <213> ORGANISM: artificial Sequence
 84 <220> FEATURE:
 85 <223> OTHER INFORMATION: Oligonucleotide
 87 <400> SEQUENCE: 5
 88 aaggtaccaa gatggactcc ccagggaagc aggatact                            38
 91 <210> SEQ ID NO: 6
 92 <211> LENGTH: 32
 93 <212> TYPE: DNA
 94 <213> ORGANISM: artificial Sequence
 96 <220> FEATURE:
 97 <223> OTHER INFORMATION: Oligonucleotide
 99 <400> SEQUENCE: 6
100 aaggatccaa gaaagttggg tggttgcagc tc                                  32
103 <210> SEQ ID NO: 7
104 <211> LENGTH: 20
105 <212> TYPE: DNA
106 <213> ORGANISM: artificial Sequence
108 <220> FEATURE:
109 <223> OTHER INFORMATION: Oligonucleotide
111 <400> SEQUENCE: 7
112 gaaagtcgga ggttcgaaga                                               20
115 <210> SEQ ID NO: 8
116 <211> LENGTH: 20
117 <212> TYPE: DNA
118 <213> ORGANISM: artificial Sequence
120 <220> FEATURE:
121 <223> OTHER INFORMATION: Oligonucleotide
123 <400> SEQUENCE: 8
124 accaactaag aacggccatg                                               20
127 <210> SEQ ID NO: 9
128 <211> LENGTH: 24
129 <212> TYPE: DNA
130 <213> ORGANISM: artificial Sequence
132 <220> FEATURE:
133 <223> OTHER INFORMATION: Oligonucleotide
135 <400> SEQUENCE: 9
136 agcaacatgc cgtcgaaccg ctcc                                          24
139 <210> SEQ ID NO: 10
140 <211> LENGTH: 24
141 <212> TYPE: DNA
142 <213> ORGANISM: artificial Sequence
144 <220> FEATURE:
145 <223> OTHER INFORMATION: Oligonucleotide
```

RAW SEQUENCE LISTING DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

<400> SEQUENCE: 10
ggagcggttc gacggcatgt tgct 24
<210> SEQ ID NO: 11
<211> LENGTH: 1281
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 11
ctagagtgca gagctccggg acgtggatcg gagccggcgc gatgggcgga gagcaggagg 60
aggagcggtt cgacggcatg ttgctggcca tggctcagca gcacgagggc ggcgtgcagg 120
agcttgtgaa caccttcttc agcttccttc gacgcaaaac agacttttc attggaggag 180
aagaagggat ggcagagaag cttatcacac agactttcag ccaccacaat cagctggcac 240
agaagacccg gcgggagaag agagcccggc aggaggccga gcggcgggag aaggcggagc 300
gggcggccag actggccaag gaagccaagt cagagacctc agggccccag atcaaggagc 360
taactgatga agaggcagag aggctgcagc tagagattga ccagaaaaag gatgcagaga 420
atcatgaggc ccagctcaag aacggcagcc ttgactcccc agggaagcag gatactgagg 480
aagatgagga ggaagatgag aaggacaaag gaaaactgaa gcccaaccta ggcaacgggg 540
cagacctgcc caattaccgc tggacccaga ccctgtcgga gctggacctg gcggtccctt 600
tctgtgtgaa cttccggctg aaagggaagg acatggtggt ggacatccag cggcggcacc 660
tccgggtggg gctcaagggg cagccagcga tcattgatgg ggagctctac aatgaagtga 720
aggtggagga gagctcgtgg ctcattgagg acggcaaggt ggtgactgtg catctggaga 780
agatcaataa gatggagtgg tggagccgct tggtgtccag tgaccctgag atcaacacca 840
agaagattaa ccctgagaat tccaagctgt cagacctgga cagtgagact cgcagcatgg 900
tggaaaagat gatgtatgac cagcgacaga agtccatggg gctgccaact tcagacgaac 960
agaagaaaca ggagattctg aagaagttca tggatcaaca tccggagatg gatttttcca 1020
aggctaaatt caactagccc ctgtttttc ctccctgaac tcttggggct gagctgcaac 1080
cacccaactt tctttcccac tcttctctgg gacttgtggg cctcagggct tggggcaggc 1140
atgggactgg cccaggcaca caggtcccgg ggcatcagga gaaaggctgg gtcttgggac 1200
cttgtcctcc ccagttggcc tactgttaca cattaaaacg atttgcccag ctcaaaaaaa 1260
aaaaaaaaaa aaaaaaaaaa a 1281
<210> SEQ ID NO: 12
<211> LENGTH: 331
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 12
Met Gly Gly Glu Gln Glu Glu Glu Arg Phe Asp Gly Met Leu Leu Ala
1 5 10 15
Met Ala Gln Gln His Glu Gly Gly Val Gln Glu Leu Val Asn Thr Phe
20 25 30
Phe Ser Phe Leu Arg Arg Lys Thr Asp Phe Phe Ile Gly Gly Glu Glu
35 40 45
Gly Met Ala Glu Lys Leu Ile Thr Gln Thr Phe Ser His His Asn Gln
50 55 60
Leu Ala Gln Lys Thr Arg Arg Glu Lys Arg Ala Arg Gln Glu Ala Glu
65 70 75 80
Arg Arg Glu Lys Ala Glu Arg Ala Ala Arg Leu Ala Lys Glu Ala Lys
85 90 95
Ser Glu Thr Ser Gly Pro Gln Ile Lys Glu Leu Thr Asp Glu Glu Ala
100 105 110
Glu Arg Leu Gln Leu Glu Ile Asp Gln Lys Lys Asp Ala Glu Asn His

RAW SEQUENCE LISTING DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

```
238         115                 120                 125
241 Glu Ala Gln Leu Lys Asn Gly Ser Leu Asp Ser Pro Gly Lys Gln Asp
242     130                 135                 140
245 Thr Glu Glu Asp Glu Glu Glu Asp Glu Lys Asp Lys Gly Lys Leu Lys
246 145                 150                 155                 160
249 Pro Asn Leu Gly Asn Gly Ala Asp Leu Pro Asn Tyr Arg Trp Thr Gln
250                 165                 170                 175
253 Thr Leu Ser Glu Leu Asp Leu Ala Val Pro Phe Cys Val Asn Phe Arg
254             180                 185                 190
257 Leu Lys Gly Lys Asp Val Val Val Asp Ile Gln Arg Arg His Leu Arg
258         195                 200                 205
261 Val Gly Leu Lys Gly Gln Pro Ala Ile Ile Asp Gly Glu Leu Tyr Asn
262     210                 215                 220
265 Glu Val Lys Val Glu Glu Ser Ser Trp Leu Ile Glu Asp Gly Lys Val
266 225                 230                 235                 240
269 Val Thr Val His Leu Glu Lys Ile Asn Lys Met Glu Trp Trp Ser Arg
270                 245                 250                 255
273 Leu Val Ser Ser Asp Pro Glu Ile Asn Thr Lys Lys Ile Asn Pro Glu
274             260                 265                 270
277 Asn Ser Lys Leu Ser Asp Leu Asp Ser Glu Thr Glu Ser Met Val Glu
278         275                 280                 285
281 Lys Met Met Tyr Asp Gln Arg Gln Lys Ser Met Gly Leu Pro Thr Ser
282     290                 295                 300
285 Asp Glu Gln Lys Lys Gln Glu Ile Leu Lys Lys Phe Met Asp Gln His
286 305                 310                 315                 320
289 Pro Glu Met Asp Phe Ser Lys Ala Lys Phe Asn
290                 325                 330
293 <210> SEQ ID NO: 13
294 <211> LENGTH: 332
295 <212> TYPE: PRT
296 <213> ORGANISM: Rattus rattus
298 <400> SEQUENCE: 13
300 Met Gly Gly Glu Gln Glu Glu Glu Arg Phe Asp Gly Met Leu Leu Ala
301 1               5                   10                  15
304 Met Ala Gln Gln His Glu Gly Gly Val Gln Glu Leu Val Asn Thr Phe
305             20                  25                  30
308 Phe Ser Phe Leu Arg Arg Lys Thr Asp Phe Phe Ile Gly Gly Glu Glu
309         35                  40                  45
312 Gly Met Ala Glu Lys Leu Ile Thr Gln Thr Phe Asn His His Asn Gln
313     50                  55                  60
316 Leu Ala Gln Lys Ala Arg Arg Glu Lys Arg Ala Arg Gln Leu Thr Glu
317 65                  70                  75                  80
320 Arg Arg Glu Lys Ala Glu Arg Ala Ala Arg Leu Ala Lys Glu Ala Lys
321                 85                  90                  95
324 Ala Glu Thr Pro Gly Pro Gln Ile Lys Glu Leu Thr Asp Leu Lys Ala
325             100                 105                 110
328 Glu Arg Leu Gln Leu Glu Ile Asp Gln Lys Lys Asp Ala Glu Asn His
329         115                 120                 125
332 Glu Val Gln Leu Lys Asn Gly Ser Leu Asp Ser Pro Gly Lys Gln Asp
```

RAW SEQUENCE LISTING DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

```
333     130                 135                 140
336 Ala Leu Leu Glu Glu Asp Glu Glu Asp Glu Lys Asp Lys Gly Lys Leu
337 145                 150                 155                 160
340 Lys Pro Asn Leu Gly Asn Gly Ala Asp Leu Pro Asn Tyr Arg Trp Thr
341                 165                 170                 175
344 Gln Thr Leu Ser Phe Leu Asp Leu Ala Val Pro Phe Arg Val Ser Phe
345             180                 185                 190
348 Arg Leu Lys Gly Lys Gln Val Val Val Asp Ile Gln Arg Arg His Leu
349         195                 200                 205
352 Arg Val Gly Leu Lys Gly Gln Ala Pro Val Ile Asp Gly Glu Leu Tyr
353     210                 215                 220
356 Asn Glu Val Lys Val Glu Glu Ser Ser Trp Leu Ile Glu Asp Gly Lys
357 225                 230                 235                 240
360 Val Val Thr Val His Leu Glu Lys Ile Asn Lys Met Glu Trp Trp Asn
361                 245                 250                 255
364 Arg Leu Val Thr Ser Asp Pro Glu Ile Asn Thr Lys Lys Ile Asn Pro
365             260                 265                 270
368 Glu Asn Ser Lys Leu Ser Asp Leu Asp Ser Glu Thr Arg Ser Met Val
369         275                 280                 285
372 Glu Lys Met Met Tyr Asp Gln Arg Gln Lys Ser Met Gly Leu Pro Thr
373     290                 295                 300
376 Ser Asp Glu Gln Lys Lys Gln Glu Ile Leu Lys Lys Phe Met Asp Gln
377 305                 310                 315                 320
380 His Pro Glu Met Asp Phe Ser Lys Ala Lys Phe Asn
381                 325                 330
384 <210> SEQ ID NO: 14
385 <211> LENGTH: 202
386 <212> TYPE: PRT
387 <213> ORGANISM: Aspergillus nidulans
389 <400> SEQUENCE: 14
391 Met Ser Glu Gln Glu Pro Ser Ser Ala Asp Leu Ala Ala Arg Glu Ala
392 1                 5                   10                  15
395 Glu Glu Lys Gln Arg Lys Ala Ala Glu Glu Ala Glu Gln Ala Thr Leu
396             20                  25                  30
399 Pro Tyr Lys Asn Thr Gln Thr Ile Arg Asp Val Asp Val Phe Thr Thr
400         35                  40                  45
403 Ile Pro Val Ser Ala Asn Leu Lys Gly Arg Asp Leu Asp Val Val Leu
404     50                  55                  60
407 Lys Lys Asp Ser Ile Lys Val Lys Val Lys Gly Glu Asn Gly Glu Val
408 65                  70                  75                  80
411 Phe Ile Asp Gly Gln Phe Pro His Pro Ile Lys Pro Ser Glu Ser Ser
412                 85                  90                  95
415 Trp Thr Leu Glu Thr Thr Ser Lys Pro Pro Phe Thr Gly Lys Glu Val
416             100                 105                 110
419 Ser Ile His Leu Asp Lys Val Asn Gln Met Glu Trp Trp Ala Met Val
420         115                 120                 125
423 Val Thr Thr Ala Pro Lys Ile Asp Val Ser Lys Ile Thr Phe Glu Asn
424     130                 135                 140
427 Ser Ser Leu Ser Asp Leu Asp Gly Glu Thr Arg Ala Met Val Glu Lys
```

FYI

Use of n and/or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY DATE: 01/15/2002
PATENT APPLICATION: US/09/623,568 TIME: 13:19:50

Input Set : A:\ES.txt
Output Set: N:\CRF3\01152002\I623568.raw

L:496 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16